Title: Robust estimation and its application to a classification problem

Author: Henryk Gacki, Agnieszka Kulawik

Citation style: Gacki Henryk, Kulawik Agnieszka. (2019). Robust estimation and its application to a classification problem. "Matematyka Stosowana" Nr 47 (2019), s. 81-94. DOI 10.14708/ma.v47i2.6499

MATHEMATICA APPLICANDA
Vol. 47(2) 2019, p. 81–91
doi: 10.14708/ma.v47i2.6480

Henryk Gacki * (Katowice)
Agnieszka Kulawik * (Katowice)

Robust estimation and its application to a classification problem

Abstract In the article a classification problem with two normally distributed classes is considered. The problem is solved using empirical discriminant functions for a Gaussian classifier and estimators for unknown parameters of the multivariate normal distribution. The following estimators will be considered: the maximum likelihood estimator (MLE), the Kulawik-Zontek estimator (KZE) and the minimum covariance determinant estimator (MCDE). Classifiers based on MLE and KZE will be compared in case of an empirical example (small sample). For large sample classifiers based on MLE, KZE and MCDE will be used.

2010 Mathematics Subject Classification: Primary: 62C12, Secondary: 62P30.

Key words and phrases: Gaussian classifier, Huber's function, estimator, multivariate normal model.

1. Introduction The main aim of classification ([18], [5]) is to make a decision which class should be attributed to a new observation. For this purpose it is possible to use a classifier based on a training set of elements whose categories (class labels) are known. In the paper a classifier based on an estimator for parameters of the multivariate normal model is considered. Hence, the classes will be assumed to be multivariate normally distributed and then a Gaussian classifier will be used. In the article two multivariate normally distributed classes are considered. Parameters - shift and positive definite covariance matrix, are unknown. Estimators of the parameters appear in the form of the empirical discriminant functions for Gaussian classifiers. In the article three estimators will be used: the maximum likelihood estimator ([15], [17]) (MLE), the Kulawik-Zontek estimator ([12]) (KZE) and the minimum covariance determinant estimator ([16]) (MCDE).

It is well known that in the case of model data - normally distributed, MLE is the best choice for estimating the unknown shift parameter and covariance matrix of the multivariate normal distribution ([18]). The situation is changing for data that is not normally distributed and it can happen for

* 0000-0002-0181-1185

* 0000-0002-2188-1295

samples with small size. For the data it is better to use robust estimators ([14]). In the paper two such estimators are considered: KZE and MCDE. The problem of the robust classification is known in literature (for example [19]). The authors compare various robust estimators. However, KZE has not been considered anywhere. The main aim of the article is to find advantages of using KZE compared to the other two estimators. In the case of a small sample it will turn out that for KZE the percentage of wrongly classified elements is lower than the one for MLE. For a large sample the percentage will be the smallest for MLE. On one hand, the percentage for KZE will be only little greater than the one for MLE but on the other hand the percentage for KZE will be clearly lower than the one for MCDE.

In the next chapter some basic information concerning Gaussian classifiers is given. In the third chapter three estimators of parameters for the multivariate normal distribution are described. In the last chapter two empirical examples are presented. The first example concerns motors (a small sample) and two estimators: the maximum likelihood estimator and the Kulawik-Zontek estimator are considered. The second example concerns the chemical analysis of wine (a large sample) and three estimators are involved: the maximum likelihood estimator, the Kulawik-Zontek estimator and the minimum covariance determinant estimator. In calculation software environment "R" was used. More precisely,

- "R" package "expm" ([6]) was used in computing KZE;
- "R" package "DetMCD" ([10]) was used in computing MCDE;
- "R" package "conics" ([3]) was used to plot the separating surfaces;
- "R" package "mvnormtest" ([9]) was used to test normality of multivariate data.

2. Gaussian classifier We consider the classification with two classes. Assume that each class is multivariate normally $N_J(\mu_i, \Sigma_i)$ distributed with the density function of the form

$$f(z|i) = (2\pi)^{-\frac{J}{2}} |\Sigma_i|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(z-\mu_i)^T \Sigma_i^{-1}(z-\mu_i)\right),$$

$z \in \mathbb{R}^J$, where μ_i, Σ_i are unknown parameters, $i = 1, 2$. Discriminant functions for the Gaussian classifier in the case of considered problem can be defined as

$$g_i(z) = \ln f(z|i) + \ln P(i),$$

for $z \in \mathbb{R}^J$, where $P(i)$ denotes "a priori probability" for the i-th class, $i = 1, 2$. For the Gaussian classifier the functions can be written in the following form

$$g_i^{\text{gauss}}(z) = -\frac{1}{2}(z-\mu_i)^T\Sigma_i^{-1}(z-\mu_i) - \frac{1}{2}\ln|\Sigma_i| + \ln P(i), \quad i = 1, 2.$$

After changing the unknown parameters $\mu_i, \Sigma_i, P(i)$ to their estimators $\hat{\mu_i}, \hat{\Sigma_i}, \hat{P(i)}$ we get the empirical discriminant functions for the Gaussian classifier:

$$\hat{g}_i^{\text{gauss}}(z) = -\frac{1}{2}(z-\hat{\mu_i})^T\hat{\Sigma_i}^{-1}(z-\hat{\mu_i}) - \frac{1}{2}\ln|\hat{\Sigma_i}| + \ln\hat{P(i)}, \tag{1}$$

$i = 1, 2$, and the formula for separating surface

$$\hat{g}_1^{\text{gauss}}(z) - \hat{g}_2^{\text{gauss}}(z) = 0 \tag{2}$$

for $z \in \mathbb{R}^J$ which gives the following decision possibility: the label $i = 1$ is assigned to an element $\hat{z}$ which is to classify, if

$$\hat{g}_1^{\text{gauss}}(\hat{z}) - \hat{g}_2^{\text{gauss}}(\hat{z}) > 0$$

and the label $i = 2$ is assigned to the element $\hat{z}$, if

$$\hat{g}_1^{\text{gauss}}(\hat{z}) - \hat{g}_2^{\text{gauss}}(\hat{z}) < 0$$

according to a maximum a posteriori probability (MAP) rule. Assume that $J = 2$ and $z = (x, y)^T \in \mathbb{R}^2$. Then,

$$\hat{\mu_1} = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}, \quad \hat{\Sigma_1}^{-1} = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix}$$

for some $m_1, m_2, \sigma_{11}, \sigma_{12}, \sigma_{22} \in \mathbb{R}$. Let $c_1 = -\frac{1}{2}\ln|\hat{\Sigma_1}| + \ln\hat{P(1)}$. We have $c_1 \in \mathbb{R}$ and

$$\hat{g}_1^{\text{gauss}}(z) = \hat{g}_1^{\text{gauss}}(\begin{bmatrix} x \\ y \end{bmatrix}) = -\frac{1}{2}\begin{bmatrix} x-m_1 \\ y-m_2 \end{bmatrix}^T \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} \begin{bmatrix} x-m_1 \\ y-m_2 \end{bmatrix} + c_1.$$

We can get an analogous form for $\hat{g}_2^{\text{gauss}}(z)$. It is clear then that the formula (2) can be written as

$$ax^2 + by^2 + cxy + dx + ey + f = 0,$$

where a, b, c, d, e, f are real coefficients.

We can use various estimators for estimating parameters μ and Σ of multivariate normal distribution $N_J(\mu, \Sigma)$. Consequently, various classifiers will be obtained.

3. Estimators for parameters of multivariate normal model

The maximum likelihood estimator (MLE) is the best choice for model data. For contaminated data it is better to use robust estimators. We will focus on two robust estimators: the Kulawik-Zontek estimator (KZE) and the minimum covariance determinant estimator (MCDE). In article [12] described results of the computer simulation are related to estimators MLE and KZE. In particular it has been shown that KZE gives better estimates than MLE for contaminated data. In our article the results will be analogous in the case of corresponding classifiers for empirical data. MCDE is one of the most popular robust estimators used by researchers. That is why we will give only a short explanation concerning MCDE (see [16]).

The maximum likelihood estimator (MLE)

Let $z_1 \in \mathbb{R}^J, \ldots, z_n \in \mathbb{R}^J$ be a value of a random sample of size n drawn from the distribution $N_J(\mu, \Sigma)$, where μ, Σ are unknown parameters ($\mu \in \mathbb{R}^J$ is an expected value and Σ is a positive definite covariance matrix). For $z_1, \ldots, z_n$ the maximum likelihood estimator (MLE) of μ and Σ is given by the formula

$$\mathrm{MLE}(\mu) = \bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i$$

and

$$\mathrm{MLE}(\Sigma) = \frac{1}{n}\sum_{i=1}^{n}(z_i - \bar{z})(z_i - \bar{z})^T,$$

respectively. When $z_1 \in \mathbb{R}^J, \ldots, z_n \in \mathbb{R}^J$ are not drawn from normal distribution, maximum likelihood estimation can give wrong results. In the case of observations it is better to use robust estimators ([14]). We will focus on two robust estimators: the Kulawik-Zontek estimator and the minimum covariance determinant estimator.

The Kulawik-Zontek estimator

The Kulawik-Zontek estimator has been described in [12]. To estimate parameters μ and Σ of the multivariate normal model the covariance matrix should be written in the form

$$\Sigma = \sum_{i=1}^{k} \alpha_i W_i,$$

where $W_1, \ldots, W_k$, $k = \frac{J(J+1)}{2}$, are the elements of a given basis of the vector space of real, square and symmetric matrices. The aim is to estimate the parameter $\theta = (\mu^T, \alpha_1, \dots, \alpha_k)^T \in \Theta \subset \mathbb{R}^{J+k}$. For the sample $z_1, \dots, z_n \in \mathbb{R}^J$ the estimator is given by the formula

$$\hat{\theta} = \operatorname{argmin}_{\theta \in \Theta} \sum_{i=1}^{n} \left\{ \ln \left| \sum_{i=1}^{k} \alpha_i W_i \right|^{\frac{1}{2}} + \varphi \left[\frac{1}{c^2} (z_i - \mu)^T \left(\sum_{i=1}^{k} \alpha_i W_i \right)^{-1} (z_i - \mu) \right] \right\} \tag{3}$$

where c is a properly chosen constant and $\varphi\colon [0, +\infty) \to \mathbb{R}$ is a function with the following properties:

(B1) The function φ has a positive derivative on $(0, +\infty)$.

(B2) The function $x\varphi'(x^2)$ has the nonnegative derivative on $[0, +\infty)$ and there exists $x_0 > 0$ such that $2x_0^2\varphi'(x_0^2) > J$.

(B3) The function φ'' is continuous.

(B4) The functions $x\varphi'(x^2)$ and $x^2\varphi''(x^2)$ are bounded.

The constant c can be derived from the equation

$$2\mathbb{E}\left(\frac{Y^T Y}{c^2} \varphi' \left(\frac{Y^T Y}{c^2} \right) \right) = J,$$

where Y is a random vector with distribution $\mathrm{N}_J(0_J, I_J)$ (multivariate standard normal distribution).

For a sample the formula (3) gives estimates $\hat{\mu}, \hat{\alpha_1}, \dots, \hat{\alpha_k}$. The covariance matrix Σ is estimated by

$$\hat{\Sigma} = \sum_{i=1}^{k} \hat{\alpha}_i W_i,$$

when it is a positive definite matrix. If not, it is possible to take $\hat{\Sigma}$ equal to square root ("R" package "expm" - function *sqrtm*) of the matrix

$$(\sum_{i=1}^{k} \hat{\alpha}_i W_i)(\sum_{i=1}^{k} \hat{\alpha}_i W_i).$$

Example 3.1 (An example of the function φ) The function $\varphi\colon [0, +\infty) \to \mathbb{R}$ given by

$$\varphi(x) = \phi(\sqrt{x}), \quad x \geq 0,$$

where the function ϕ is defined by its derivative

$$\phi'(x) = \begin{cases} x, & |x| \leq t, \\ -x - 4t - \frac{2t^2}{x}, & -2t < x < -t, \\ -x + 4t - \frac{2t^2}{x}, & t < x < 2t, \\ \frac{2t^2}{x}, & |x| \geq 2t. \end{cases}$$

and $t > 0$, satisfies the conditions (B1)-(B4).

The above functions are modifications of Huber functions ([7]). The family from example 3.1 consists of functions that depend on the tunning constant t. The tunning constant is usually given. However, for some models it is possible to get the data dependent tuning constant (for example [13]). In the case of our article modifications of the functions will be used with a proper given constant. This type of modifications has already been used by T. Bednarski and S. Zontek in [2].

The minimum covariance determinant estimator (MCDE)

The minimum covariance determinant estimator ([16]) is a robust estimator of the expectation and covariance matrix for the multivariate normal distribution. The estimator is based on the subset of all given observations for which the covariance matrix has the smallest determinant. The mean of the elements from the chosen subset is the minimum covariance determinant estimator (MCDE) of the population mean and their covariance matrix is MCDE of the population covariance matrix. For empirical problems with multivariate data it is better to use an approximation of MCDE's values rather than the exact ones because of computation time. To get an approximation the Deterministic Minimum Covariance Determinant algorithm ([8]) can be used.

4. Empirical examples

Motor problem

In article [1] the authors have presented an attempt to draw the image of the cognition method application in the diagnostic experiment carried out in tests of the single-phase induction motors. More precisely, motors of type SZXb6514 B made by Zakład Silników Elektrycznych Małej Mocy "Silma" in Sosnowiec (Low Power Electric Motors Company "Silma") were considered (11 usable motors and 23 that are not usable). The motors were represented by 9-dimensional vectors and the authors used the Karhunen-Loeve (K-L) method ([5]) and got 2-dimensional vectors. Not usable motors were grouped with respect to the following types of defects:

- B - rubbing,
- C - loudness - loud operation,
- E - high current,
- F - increased vibration level,
- G - no rivet in the sheet package.

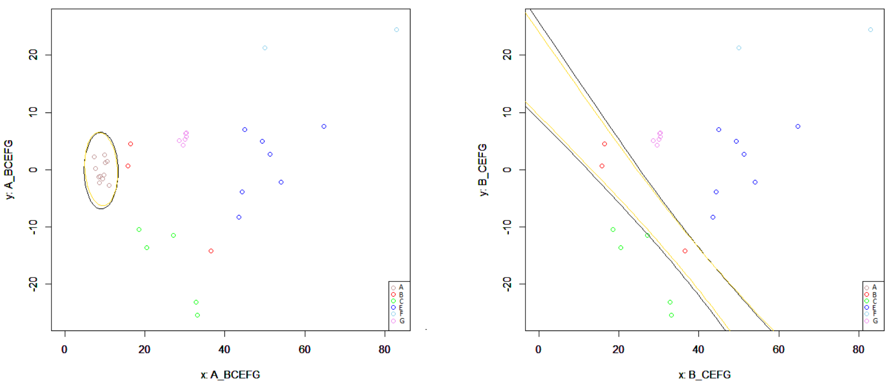
The following classification problems with two classes were considered:

- usable motors (A) - not usable motors (BCEFG),
- B - CEFG,
- C - BEFG,
- E - BCFG,
- F - BCEG,
- G - BCEF.

Gaussian classifiers using the empirical discriminant functions for the estimators

- MLE
- KZE (for function φ from Example 3.1 with $t = 1,445$ and $c = 0,865$, for which the loss of efficiency of estimating the shift parameter equals 10% according to MLE)

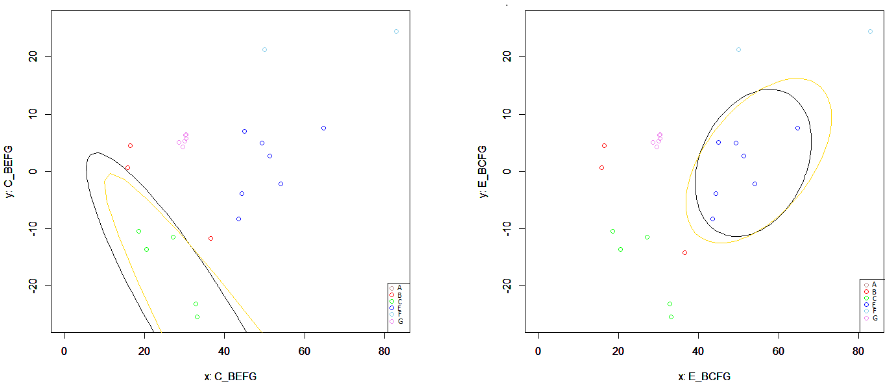
will be compared. Figures 1, 2 and 3 present the image of the 9-dimensional vectors (primary features) transformed to 2-dimensional vectors (secondary features) using the K-L method. The figures present also separating surfaces for the considered cases. Orange is used for MLE and black is for KZE. The case F - BCEG is omitted because there were only two motors with F defect.



(a) A-BCEFG (b) B-CEFG

Figure 1: Separating surfaces

Fractions of wrongly classified elements (Leave-One-Out Method) are given in Table 1. For the case E - BCFG the fraction is smaller than the one for MLE. In other cases the fractions are the same. An analysis concerning normality of the proper datasets was done. Shapiro-Wilk normality test for multivariate data allowed us to reject the hypothesis of normality for the sets: B, C, G, BCEFG, BEFG, BCFG (red in Table 1).



(a) C-BEFG (b) E-BCFG

Figure 2: Separating surfaces

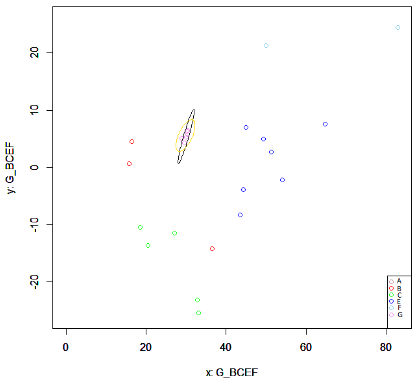


Figure 3: Separating surfaces

Table 1: Fractions of wrongly classified elements

	MLE	KZE
A - BCEFG	0	0
B - CEFG	0	0
C - BEFG	1/23	1/23
E - BCFG	4/23	3/23
G - BCEF	1/23	1/23

The percentages of errors (wrongly classified elements) are given in Table 2. The percentage for KZE (4,35%) is lower than the one for MLE (5,22%). The results show that for the small size sample problem it was better to use

KZE than MLE.

Table 2: Percentages of wrongly classified elements

MLE	KZE
$5,22\%$	$4,35\%$

Wine problem

The dataset WINE has been used:
source: D. Dua and E. Karra Taniskidou (2017). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science ([4]).
The database presents the effect of three types (type 1, type 2, type 3) of cultivars on the chemical analysis of wines from the same region in Italy. The following chemical features are investigated: alcohol, malic acid, ash, alcalinity of ash, magnesium, total phenols, flavanoids, nonflavanoid phenols, proanthocyanins, color intensity, hue, OD280/OD315 of diluted wines, proline. The analysis based on 144 13-dimensional vectors taken from the dataset WINE (48 vectors for each type of cultivars). Using the linear discriminant analysis ([11]) 2-dimensional vectors have been obtained ("R" package "MASS" - function *lda*). We are interested in the classification with two classes, hence the data was divided into three problems:

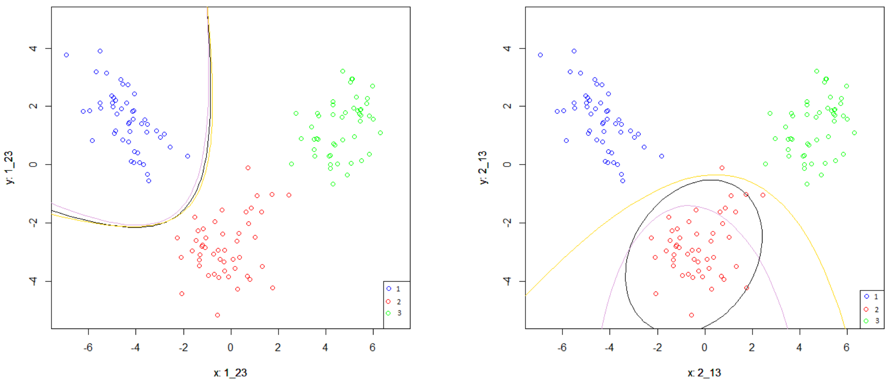
- type 1 - types 2,3,
- type 2 - types 1,3,
- type 3 - types 1,2.

In the problems Gaussian classifiers using the empirical discriminant functions for estimators:

- MLE,
- KZE (with $t = 1,445$ and $c = 0,865$),
- MCDE,

will be compared. Figures 4 and 5 present the image of the 13-dimensional vectors (primary features) transformed to 2-dimensional vectors (secondary features) using the linear discriminant analysis. The figures present also separating surfaces for the considered cases. Orange is used for MLE, black is for KZE and purple is for MCDE.

Fractions of the wrongly classified elements (10-fold cross-validation) are given in Table 3. The percentage of errors (the wrongly classified elements) are given in Table 4. In the case 1-(2,3) (Table 3) all estimators classified



(a) 1-23 (b) 2-13

Figure 4: Separating surfaces

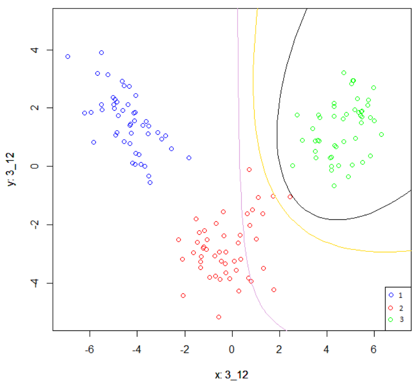


Figure 5: Separating surfaces

elements correctly. For 2-(1,3) two robust estimators were worse than MLE. In the case 3-(1,2) for MCDE we notice the highest number of the wrongly classified elements. Meanwhile for KZE the number is zero. The percentages of the wrongly classified elements show that the best choice for the problem is MLE but also KZE. MCDE gives the worst results. An analysis concerning normality of the proper datasets was done. Shapiro-Wilk normality test for multivariate data did not allow us to reject the hypothesis of normality in any case.

Conclusions

KZE is a modification of MLE. Using an empirical example (the motor problem) we compared the corresponding two classifiers. The percentage of

Table 3: Fractions of wrongly classified elements

	MLE	KZE	MCDE
1 - (2,3)	0	0	0
2 - (1,3)	1/144	4/144	4/144
3 - (1,2)	2/144	0	5/144

Table 4: Percentages of wrongly classified elements

MLE	KZE	MCDE
0,69%	0,93%	2,08%

the wrongly classified elements is lower in the case of the classifier using the KZE estimator. KZE is a robust estimator, so another robust estimator was chosen for further analysis - one of the most popular (MCDE). We compared the three classifiers for large samples (normal samples). It occurs that the classifier using KZE is not much worse than the classifier using MLE. Meanwhile, classifier for MCDE seems to give much worse results.

To sum up, we can see that in the case of the real data (the motors problem) for which the hypothesis of normality is rejected the classifiers based on robust estimators are a better choice than the classifier based on MLE. However, for a sample for which the hypothesis concerning normality is not rejected (wine problem) the classifier based on robust KZE seems to be better than the classifier based on robust MCDE.

Conducting research on real data has shown that KZE can be used in industry. In the case of engines KZE allows to characterize and distinguish types of damage more precisely in comparison with MLE. This can be seen for example in Figure 2(b) and Figure 3.

Acknowledgments

We want to thank prof. dr hab. Mirosław Krzyśko for important remarks concerning the linear discriminant analysis. The remarks proved to be useful in getting the interesting results of the wine analysis.

References

[1] J. Adamczewski and H. Gacki. Próba zastosowania metod rozpoznawania obrazów w eksperymencie diagnostycznym. *Zagadnienia eksploatacji maszyn*, 97(1):161–172, 1994. Cited on p. 86.

[2] T. Bednarski and S. Zontek. Robust estimation of parameters in a mixed unbalanced model. *Ann. Statist.*, 24(4):1493–1510, 1996. Cited on p. 86.

[3] B. Desgraupes. conics: Plot conics. *"R" package version 0.3*, 2013. Cited on p. 82.

[4] D. Dua and E. K. Taniskidou. UCI Machine Learning Repository. *University of California, Irvine, School of Information and Computer Sciences*, 2017. URL http://archive.ics.uci.edu/ml. Cited on p. 89.

[5] R. C. Gonzalez and J. T. Tou. *Pattern recognition principles*. Addison-Wesley Publ. Comp., 1974. Cited on pp. 81 and 86.

[6] V. Goulet, C. Dutang, M. Maechler, D. Firth, M. Shapira, and M. Stadelmann. expm: Matrix exponential, log, 'etc'. *"R" package version 0.999-4*, 2019. Cited on p. 82.

[7] P. J. Huber. *Robust Statistics*. Wiley, New York, 1981. Cited on p. 86.

[8] M. Hubert, P. J. Rousseeuw, and T. Verdonck. A deterministic algorithm for robust location and scatter. *Journal of Computational and Graphical Statistics*, 21(3):618–637, 2012. Cited on p. 86.

[9] S. Jarek. mvnormtest: Normality test for multivariate variables. *"R" package version 0.1-9*, 2015. Cited on p. 82.

[10] V. Kaveh and M. Hubert. DetMCD: Implementation of the DetMCD algorithm (robust and deterministic estimation of location and scatter). *"R" package version 0.0.5*, 2018. Cited on p. 82.

[11] M. Krzyśko. *Podstawy wielowymiarowego wnioskowania statystycznego*. Wydawnictwo Naukowe UAM, Poznań, 2009. Cited on p. 89.

[12] A. Kulawik and S. Zontek. Robust estimation in the multivariate normal model. *Discussiones Mathematicae Probability and Statistics*, 36(1-2): 53–66, 2016. Cited on pp. 81 and 84.

[13] A. Kulawik and S. Zontek. On data-dependent tuning constant of the huber family for robust estimator. *(to be published)*, 2018. Cited on p. 86.

[14] R. Maronna, D. Martin, and V. Yohai. *Robust Statistics: Theory and Methods*. John Wiley and Sons, 2006. Cited on pp. 82 and 84.

[15] M. Nieniewski. *Probabilistic methods in pattern recognition and computer vision*. Lecture Notes and Computer Vision and Artificial Intelligence, Ossolineum, 1990. Cited on p. 81.

[16] P. J. Rousseeuw. Multivariate estimation with high breakdown point. *Mathematical Statistics and Applications*, Vol. B (Bad Tatzmannsdorf, 1983), Reidel, Dordrecht(1-2):283–297, 1985. Cited on pp. 81, 84, and 86.

[17] W. Sobczak and W. Malina. *Metody selekcji i redukcji informacji*. WNT, Warszawa, 1985. Cited on p. 81.

[18] N. H. Timm. *Applied Multivariate Analysis*. Springer-Verlag New York, 2002. Cited on p. 81.

[19] V. Todorov and A. M. Pires. Comparative performance of several robust linear discriminant analysis methods. *REVSTAT-Statistical Journal*, 5 (1):63–83, 2007. Cited on p. 82.

Estymacja odporna i jej zastosowanie w pewnym problemie klasyfikacji

Henryk Gacki, Agnieszka Kulawik

Streszczenie W artykule omówiono problem klasyfikacji dla dwóch klas w przypadku przyjęcia założenia, że rozkłady cech w klasach są wielowymiarowymi rozkładami normalnymi. Problem rozwiązano za pomocą empirycznego klasyfikatora gaussowskiego i wybranych estymatorów nieznanych parametrów wielowymiarowego rozkładu normalnego. Uwzględnione zostały następujące estymatory: MLE (the maximum likelihood estimator - estymator największej wiarogodności), KZE (Kulawik-Zontek estimator) i MCDE (the minimum covariance determinant estimator). Klasyfikatory oparte o MLE i KZE zostały porównane w przypadku przykładu empirycznego (mała próba). W przypadku dużych prób porównane zostały klasyfikatory oparte o trzy wspomniane estymatory.

Klasyfikacja tematyczna AMS (2010): 62C12; 62P30.

Słowa kluczowe: klasyfikator gaussowski, funkcja Hubera, estymator, wielowymiarowy rozkład normalny.

Henryk Gacki[a] was awarded a Master of Science diploma in Mathematics in 1976, the University of Silesia, PhD degree in mathematics from the University of Silesia in 1984 and DSC (habilitation) degree in 2008 in mathematics from the University of Silesia. In 2016 Professorship at the University of Silesia in Katowice. For the year 2016 Chair of the Section of Mathematical Methods in Economy and Finances, the Institute of Mathematics, the University of Silesia. In the years 2012 – 2015 Coordinator of Mathematics as a code of Modernity European Programme under whose framework a target field of study was being realized at the University of Silesia. Since 2015 Intermaths Polish–Italian Master's Double Degree Programme Coordinator at the University of Silesia. For the year 2014 Statistical editor in "Chowanna". A member of the Polish Mathematical Association. Scientific interests: Theory of semigroups of Markov operators and their applications, testing stability of solutions of different types of the Boltzmann equation on the space of signed measures, generalizations and applications of the Kantorovich–Rubinstein maximum principle constituting an element of the solution in the transport of mass problem, probability theory and statistics. Distinctions and awards: 2016 Golden Badge of Merit for the University of Silesia, 2018 Golden Honorary Badge for Merits for the Silesian Voivodship, 2018 Selected Athlete of the 50th anniversary of AZS the University of Silesia.

[a]References to his research papers are found in MathSciNet under ID:223263 and the European Mathematical Society, FIZ Karlsruhe, and the Heidelberg Academy of Sciences bibliography database known as zbMath under ai:Gacki.Henryk.

Agnieszka Kulawik received the MSc degree in Mathematics from the University of Silesia in Katowice (2007) and PhD degree in Mathematics from the University of Zielona Góra (2011). She is interested in the following research topics: mathematical statistics and application of statistics. References to her research papers are found in MathSciNet under ID:1115942 and the European Mathematical Society, FIZ Karlsruhe, and the Heidelberg Academy of Sciences bibliography database known as zbMath under ai:Kulawik.Agnieszka.

Henryk Gacki
University of Silesia in Katowice
Institute of Mathematics
Bankowa 14, 40-007 Katowice, Poland
E-mail: henryk.gacki@us.edu.pl

Agnieszka Kulawik
University of Silesia in Katowice
Institute of Mathematics
Bankowa 14, 40-007 Katowice, Poland
E-mail: agnieszka.kulawik@us.edu.pl

Communicated by: Łukasz Stettner

(Received: 13th of April 2018; revised: 14th of June 2018)